

Experimental validation of volume-based comparison for double-McCormick relaxations*

Emily Speakman, Han Yu, and Jon Lee

Department of Industrial and Operations Engineering
University of Michigan
Ann Arbor, MI, U.S.A.

{eespeakm, yuha, jonxlee}@umich.edu

August 9, 2016

Abstract

Volume is a natural geometric measure for comparing polyhedral relaxations of non-convex sets. Speakman and Lee gave volume formulae for comparing relaxations of trilinear monomials, quantifying the strength of various natural relaxations. Their work was motivated by the spatial branch-and-bound algorithm for factorable mathematical-programming formulations. They mathematically analyzed an important choice that needs to be made whenever three or more terms are multiplied in a formulation. We experimentally substantiate the relevance of their main results to the practice of global optimization, by applying it to difficult box cubic problems (boxcup). In doing so, we find that, using their volume formulae, we can accurately predict the quality of a relaxation for boxcups based on the (box) parameters defining the feasible region.

*Supported in part by ONR grant N00014-14-1-0315.

1 Introduction

1.1 Measuring relaxations via volume in mathematical optimization

In the context of mathematical optimization, there is often a natural tradeoff in the tightness of a *convexification* (i.e., a convex relaxation) and the difficulty of optimizing over it. This idea was emphasized by [Lee07] in the context of mixed-integer nonlinear programming (MINLP) (see also the recent work [DMW15]). Of course this is also a well-known phenomenon for difficult 0/1 linear-optimization problems, where very tight relaxations are available via extremely heavy semidefinite-programming relaxations (e.g., the Lasserre hierarchy), and the most effective relaxation for branch-and-bound/cut may well not be the tightest. Earlier, again in the context of mathematical optimization, [LM94] introduced the idea of using volume as a measure of the tightness of a convex relaxation (for fixed-charge and vertex packing problem). Most of that mathematical work was asymptotic, seeking to understand the quality of families of relaxations with a growing number of variables, but some of it was also substantiated experimentally in [Lee07].

1.2 Spatial branch-and-bound

The workhorse algorithm for global optimization of so-called factorable MINLPs (see [McC76]) is spatial branch-and-bound (sBB) (see [ADFN98], [RS96], [SP99]). sBB decomposes model functions against a library of basic functions (e.g., x_1x_2 , $x_1x_2x_3$, $x_1^{x_2}$, $\log(x)$, $\sin(x)$, x^3 , \sqrt{x} , $\arctan(x)$). We assume that each basic function is a function of no more than 3 variables, and that we have a convex outer-approximation of the graph of each such basic function on box domains. sBB realizes convexifications of model functions by composing convexifications of basic functions. sBB subdivides the domains of variables, and re-convexifies, obtaining stronger bounds on model functions.

1.3 Using volume to guide decompositions for spatial branch-and-bound

[SL15] applied the idea of [LM94], but now in the context of the low-dimensional relaxations of basic functions that arise in sBB. Specifically, [SL15] considered the basic function $f = x_1x_2x_3$ on box domains; that is the graph

$$\mathcal{P}_h = \{(f, x_1, x_2, x_3) \in \mathbb{R}^4 : f = x_1x_2x_3, x_i \in [a_i, b_i], i = 1, 2, 3\},$$

where $0 \leq a_i < b_i$ are given constants. It is important to realize that \mathcal{P}_h is relevant for any model where any three quantities (which could be complicated functions themselves) are multiplied. Furthermore, the case of nonzero lower bounds (i.e., $a_i > 0$) is particularly relevant, especially when the multiplied quantities are complicated functions of model variables. Though polyhedral, \mathcal{P}_h has a rather complicated inequality description. Often lighter relaxations are used by modelers and MINLP software. [SL15] considered 3 natural relaxations of \mathcal{P}_h . Thinking of the product as $x_1(x_2x_3)$ or $x_2(x_1x_3)$ or $x_3(x_1x_2)$, and employing the so-called McCormick relaxation twice, leads to three *different* relaxations. [SL15] derive analytic expressions for the volume of \mathcal{P}_h as well as all three of the natural relaxations. The expressions are formulae in the six constants $0 \leq a_i < b_i$, $i = 1, 2, 3$. In doing so, they quantify the quality of the various relaxations and provide recommendations for which to use.

1.4 Our contribution

The results of [SL15] are theoretical. Their utility for guiding modelers and sBB implementers depends on the belief that volume is a good measure of the quality of a relaxation. Morally, this

belief is based on the idea that with no prior information on the form of an objective function, the solution of a relaxation should be assumed to occur with a uniform density on the feasible region. Our contribution is to experimentally validate the robustness of this theory in the context of a particular use case, optimizing multilinear cubics over boxes (‘boxcup’). There is considerable literature on techniques for optimizing quadratics, much of which is developed and validated in the context of so-called ‘boxqp’ problems, where we minimize $\sum_{i,j} q_{ij}x_i x_j$ over a box domain in \mathbb{R}^n . So our boxcup problems, for which we minimize $\sum_{i,j,k} q_{ijk}x_i x_j x_k$ over a box domain in \mathbb{R}^n , are natural.

[SL15] find an ordering of the three natural relaxations by volume. But their formulae are for $n = 3$. Our experiments validate their theory as applied to our use case. We demonstrate that in the setting of ‘boxcup’ problems, the average objective discrepancy between relaxations very closely follows the prediction of the theory of [SL15]. Moreover, we are able to demonstrate that these results are robust against sparsity of the cubic forms.

[LM94] defined the *idealized radius* of a polytope in \mathbb{R}^d as essentially the d -th root of its volume (up to some constants depending on d). For a polytope that is very much like a ball in shape, we can expect that this quantity is (proportional to) the “average width” of the polytope. The average width arises by looking at ‘max minus min’, averaged over all normalized linear objectives. So, the implicit prediction of [SL15] is that the *idealized radius* should (linearly) predict the expected ‘max minus min’ for normalized linear objectives. We have validated this experimentally, and looked further into the *idealized radial distance* between pairs of relaxations, finding an even higher degree of linear association.

Finally, in the important case $a_1 = a_2 = 0$, $b_1 = b_2 = 1$, [SL15] found that the two worst relaxations have the same volume, and the greatest difference in volume between \mathcal{P}_h and the (two) worst relaxations occurs when $a_3 = b_3/3$. We present results of experiments that clearly show that these predictions via volume are again borne out on ‘boxcup’ problems.

All in all, we present convincing experimental evidence that volume is a good predictor for quality of relaxation in the context of sBB. Our results strongly suggest that the theoretical results of [SL15] are important in devising decompositions of complex functions in the context of factorable formulations and therefore our results help inform both modelers and implementers of sBB.

1.5 Literature review

Computing the volume of a polytope is well known to be strongly #P-hard (see [BW91]). But in fixed dimension, or, in celebrated work, by seeking an approximation via a randomized algorithm (see [DFK91]), positive results are available. Our work though is motivated not by algorithms for volume calculation, but rather in certain situations where analytic formulae are available.

Besides [LM94] and [SL15] (and the follow-on [SL16]), there have been a few papers on analytic formulae for volumes of polytopes that naturally arise in mathematical optimization; see [KLS97], [Ste94], [BDLO13], [ABD10], [Sta86]. But none of these works has attempted to apply their ideas to the low-dimensional polytopes that naturally arise in sBB, or even to apply their ideas to compare relaxations. One notable exception is [CLL10], which is a mostly computational precursor to [SL15], focusing on quadrilinear functions (i.e., $f = x_1 x_2 x_3 x_4$).

There are many implementations of sBB. E.g., Baron [Sah15], Couenne [BLL⁺09], Scip [VG16] and Antigone [MF14]. Both Baron and Antigone use the complete linear-inequality description of \mathcal{P}_h , while Couenne and SCIP use an *arbitrary* double-McCormick relaxation. Our results indicate that there are situations where the choice of Baron and Antigone is too heavy, and certainly even restricting to double-McCormick relaxations, Couenne and SCIP do not choose the best one.

There is a large literature on convexification of multilinear functions. Most relevant to our work are: the polyhedral nature of the convexification of the graphs of multilinear functions on box

domains (see [Rik97]); the McCormick inequalities describing giving the complete linear-inequality description for bilinear functions on a box domain (see [McC76]); the complete linear-inequality description of \mathcal{P}_h (see [MF04b] and [MF04a]).

2 Preliminaries

2.1 Convexifications

Without loss of generality, we can relabel the three variables so that

$$a_1b_2b_3 + b_1a_2a_3 \leq b_1a_2b_3 + a_1b_2a_3 \leq b_1b_2a_3 + a_1a_2b_3. \quad (\Omega)$$

is satisfied by the variable bounds. Given $f := x_1x_2x_3$ satisfying (Ω) , there are three choices of double-McCormick convexifications depending on the bilinear sub-monomial we convexify first. We could first group x_1 and x_2 and convexify $w = x_1x_2$; after this, we are left with the monomial $f = wx_3$ which we again convexify using McCormick. Alternatively, we could first group variables x_1 and x_3 , or variables x_2 and x_3 .

To see how to perform these convexifications in general, we exhibit the double-McCormick convexification that first groups the variables x_i and x_j . Therefore we have $f = x_ix_jx_k$, and we let $w_{ij} = x_ix_j$, so $f = w_{ij}x_k$. Convexifying $w_{ij} = x_ix_j$ using the standard McCormick inequalities, and then convexifying $f = w_{ij}x_k$, again using the standard McCormick inequalities, we obtain the 8 inequalities:

$$\begin{aligned} w_{ij} - a_jx_i - a_ix_j + a_ia_j &\geq 0, & f - a_kw_{ij} - a_ia_jx_k + a_ia_ja_k &\geq 0, \\ -w_{ij} + b_jx_i + a_ix_j - a_ib_j &\geq 0, & -f + b_kw_{ij} + a_ia_jx_k - a_ia_jb_k &\geq 0, \\ -w_{ij} + a_jx_i + b_ix_j - b_ia_j &\geq 0, & -f + a_kw_{ij} + b_ib_jx_k - b_ib_ja_k &\geq 0, \\ w_{ij} - b_jx_i - b_ix_j + b_ib_j &\geq 0, & f - b_kw_{ij} - b_ib_jx_k + b_ib_jb_k &\geq 0. \end{aligned}$$

Using Fourier-Motzkin elimination (i.e., projection), we then eliminate the variable w_{ij} to obtain the following system in our original variables f, x_i, x_j and x_k .

$$\begin{aligned} x_i - a_i &\geq 0, & (1) \\ x_j - a_j &\geq 0, & (2) \\ f - a_ja_kx_i - a_ia_kx_j - a_ia_jx_k + 2a_ia_ja_k &\geq 0, & (3) \\ f - a_jb_kx_i - a_ib_kx_j - b_ib_jx_k + a_ia_jb_k + b_ib_jb_k &\geq 0, & (4) \\ -x_j + b_j &\geq 0, & (5) \\ -x_i + b_i &\geq 0, & (6) \\ f - b_ja_kx_i - b_ia_kx_j - a_ia_jx_k + a_ia_ja_k + b_ib_ja_k &\geq 0, & (7) \\ f - b_jb_kx_i - b_ib_kx_j - b_ib_jx_k + 2b_ib_jb_k &\geq 0, & (8) \\ -f + b_jb_kx_i + a_ib_kx_j + a_ia_jx_k - a_ia_jb_k - a_ib_jb_k &\geq 0, & (9) \\ -f + a_jb_kx_i + b_ib_kx_j + a_ia_jx_k - a_ia_jb_k - b_ia_jb_k &\geq 0, & (10) \\ -x_k + b_k &\geq 0, & (11) \\ -f + b_ja_kx_i + a_ia_kx_j + b_ib_jx_k - a_ib_ja_k - b_ib_ja_k &\geq 0, & (12) \\ -f + a_ja_kx_i + b_ia_kx_j + b_ib_jx_k - b_ia_ja_k - b_ib_ja_k &\geq 0, & (13) \\ x_k - a_k &\geq 0, & (14) \\ f - a_ia_jx_k &\geq 0, & (15) \\ -f + b_ib_jx_k &\geq 0. & (16) \end{aligned}$$

It is easy to see that the inequalities (15) and (16) are redundant: (15) is $a_j a_k(1) + a_i a_k(2) + (3)$, and (16) is $b_j a_k(6) + a_i a_k(5) + (12)$.

We use the following notation in what follows. For $i = 1, 2, 3$, *system i* is defined to be the system of inequalities obtained by first grouping the pair of variables x_j and x_k , with j and k different from i . \mathcal{P}_i is defined to be the solution set of this system.

2.2 Volumes of convexifications

For a convex body $C \subset \mathbb{R}^d$, we denote its *volume* (i.e., Lebesgue measure) by $\text{vol}(C)$. The main results of [SL15] are as follows.

Theorem 1 ([SL15])

$$\text{vol}(\mathcal{P}_h) = (b_1 - a_1)(b_2 - a_2)(b_3 - a_3) \times \\ (b_1(5b_2b_3 - a_2b_3 - b_2a_3 - 3a_2a_3) + a_1(5a_2a_3 - b_2a_3 - a_2b_3 - 3b_2b_3)) / 24.$$

Theorem 2 ([SL15])

$$\text{vol}(\mathcal{P}_1) = \text{vol}(\mathcal{P}_h) + (b_1 - a_1)(b_2 - a_2)^2(b_3 - a_3)^2 \times \\ \frac{3(b_1b_2a_3 - a_1b_2a_3 + b_1a_2b_3 - a_1a_2b_3) + 2(a_1b_2b_3 - b_1a_2a_3)}{24(b_2b_3 - a_2a_3)}.$$

Theorem 3 ([SL15])

$$\text{vol}(\mathcal{P}_2) = \text{vol}(\mathcal{P}_h) + \frac{(b_1 - a_1)(b_2 - a_2)^2(b_3 - a_3)^2 (5(a_1b_1b_3 - a_1b_1a_3) + 3(b_1^2a_3 - a_1^2b_3))}{24(b_1b_3 - a_1a_3)}.$$

Theorem 4 ([SL15])

$$\text{vol}(\mathcal{P}_3) = \text{vol}(\mathcal{P}_h) + \frac{(b_1 - a_1)(b_2 - a_2)^2(b_3 - a_3)^2 (5(a_1b_1b_2 - a_1b_1a_2) + 3(b_1^2a_2 - a_1^2b_2))}{24(b_1b_2 - a_1a_2)}.$$

Corollary 1 ([SL15])

$$\text{vol}(\mathcal{P}_h) \leq \text{vol}(\mathcal{P}_3) \leq \text{vol}(\mathcal{P}_2) \leq \text{vol}(\mathcal{P}_1).$$

Corollary 2 ([SL15]) *For the special case of $a_1 = a_2 = 0$, $b_1 = b_2 = 1$, and fixed b_3 , the greatest difference in volume for $\mathcal{P}_3 (= \mathcal{P}_h)$ and \mathcal{P}_2 (or \mathcal{P}_1) occurs when $a_3 = b_3/3$.*

2.3 From volumes to gaps

Volume seems like an awkward measure to compare relaxations, when typically we are interested in objective-function gaps. Following [LM94], the *idealized radius* of a convex body $C \subset \mathbb{R}^d$ is

$$\rho(C) := (\text{vol}(C)/\text{vol}(B_d))^{1/d},$$

where B_d is the (Euclidean) unit ball in \mathbb{R}^d . $\rho(C)$ is simply the radius of a ball having the same volume as C . The *idealized radial distance* between convex bodies C_1 and C_2 is simply $|\rho(C_1) - \rho(C_2)|$. If C_1 and C_2 are concentric balls, say with $C_1 \subset C_2$, then the idealized radial distance between them is the (radial) height of C_2 above C_1 . The *mean semi-width* of C is simply

$$\frac{1}{2} \int_{\|c\|=1} \left(\max_{x \in C} c'x - \min_{x \in C} c'x \right) d\psi,$$

where ψ is the $(d - 1)$ -dimensional Lebesgue measure on the boundary of B_d , normalized so that ψ on the entire boundary is unity. If C is itself a ball, then (i) its idealized radius is in fact its radius, and (ii) its width in any unit-norm direction c is constant, and so (iii) its (idealized) radius is equal to its mean semi-width.

Key point: What we can hope is that our relaxations are round enough so that choosing one of small volume (which is proportional to the idealized radius) is a good proxy for choosing the relaxation by *mean width* (which is the same as mean objective-value range).

3 Computational Experiments

3.1 Box cubic programs and 4 relaxations

Our experiments are aimed at the following natural problem which is concerned with optimizing a linear function on trinomials. Let H be a 3-uniform hyper-graph on n vertices. Each hyper-edge of H is a set of 3 vertices, and we denote the set of hyperedges by $E(H)$. If H is complete, then $|E(H)| = \binom{n}{3}$. We associate with each vertex i a variable $x_i \in [a_i, b_i]$, and with each hyper-edge $\{i, j, k\}$ the trinomial $x_i x_j x_k$ and a coefficient q_{ijk} ($1 \leq i < j < k \leq n$). We now formulate the associated *boxcup* ('box cubic problem'):

$$\min_{x \in \mathbb{R}^n} \left\{ \sum_{\{i,j,k\} \in E(H)} q_{ijk} x_i x_j x_k : x_i \in [a_i, b_i], i = 1, 2, \dots, n \right\}. \quad (\text{BCUP})$$

The name is in analogy with the well-known *boxqp*, where just two terms (rather than three) are multiplied ('box' refers to the feasible region and 'qp' refers to 'quadratic program').

(BCUP) is a difficult nonconvex global-optimization problem. Our goal here is not to solve instances of this problem, but rather to solve a number of different *relaxations* of the problem and see how the results of these experiments correlate with the volume results of [SL15]. In this way, we seek to determine if the guidance of [SL15] is relevant to modelers and those implementing sBB.

We have seen how for a single trilinear term $f = x_i x_j x_k$, we can build four distinct relaxations: the convex hull of the feasible points, \mathcal{P}_h , and three relaxations arising from double McCormick: \mathcal{P}_1 , \mathcal{P}_2 and \mathcal{P}_3 . To obtain a relaxation of (BCUP), we choose a relaxation \mathcal{P}_ℓ , for some $\ell = 1, 2, 3, h$ and apply this same relaxation method to *each* trinomial of (BCUP). We therefore obtain 4 distinct linear relaxations of the form:

$$\min_{(x,f) \in \mathcal{P}_\ell} \left\{ \sum_{\{i,j,k\} \in E(H)} q_{ijk} f_{ijk} \right\}.$$

where \mathcal{P}_ℓ , $\ell = 1, 2, 3, h$ is the polytope in dimension $\binom{n}{3} + n$ arising from using relaxation \mathcal{P}_ℓ on each trinomial. This linear relaxation is a linear inequality system involving the n variables x_i ($i = 1, 2, \dots, n$), and the $\binom{n}{3}$ new 'function variables' f_{ijk} . These 'function variables' model $x_i x_j x_k$.

For our experiments, we randomly generate box bounds $[a_i, b_i]$ on x_i , for each $i = 1, \dots, n$ independently, by choosing (uniformly) random pairs of integers $0 \leq a_i < b_i \leq 10$. With each realization of these bounds, we get relaxation feasible regions \mathcal{P}_ℓ , for $\ell = 1, 2, 3, h$.

3.2 3 scenarios for Q

We have 3 scenarios for the hypergraph H of (BCUP), all with $|E(H)| = 20$ monomials:

- Our **dense** scenario has H being a *complete* 3-uniform hypergraph on $n = 6$ vertices ($\binom{6}{3} = 20$). We note that each of the $n = 6$ variables appears in $\binom{6-1}{3-1} = 10$ of the 20 monomials, so there is considerable overlap in variables between trinomials.
- Our **sparse** scenario has hyperedges: $\{1, 2, 3\}, \{2, 3, 4\}, \{3, 4, 5\} \dots \{18, 19, 20\}, \{19, 20, 1\}, \{20, 1, 2\}$. Here we have $n = 20$ variables and each variable is in only 3 of the trinomials.
- Our **very sparse** scenario has $n = 30$ variables and each variable is in only 2 of the trinomials.

For each scenario, we generate 30 *sets* of bounds $[a_i, b_i]$ on x_i ($i = 1, \dots, n$). To control the variation in our results, and considering that the scaling of Q is arbitrary, we generate 100,000 random Q with $|E(H)|$ entries, uniformly distributed on the unit sphere in $\mathbb{R}^{|E(H)|}$. Then, for each Q , we both minimize and maximize $\sum_{i < j < k} q_{ijk} f_{ijk}$ over each $\mathcal{P}_\ell, \ell = 1, 2, 3, h$ and each set of bounds.

3.3 Quality of relaxations

For each Q we take the difference in the optimal values, i.e. the maximum value minus the minimum value; this can be thought of as the width of the polytope in the direction Q . We then average these widths for each $\mathcal{P}_\ell, \ell = 1, 2, 3, h$, across the 100,000 realizations of Q (which results in very small standard errors), and we refer to this quantity $\omega(\mathcal{P}_\ell)$ as the *quasi mean width* of the relaxation. It is not quite the geometric mean width, because we do not have objective terms for all variables in (BCUP) (i.e., we have no objective terms $\sum_{i=1}^n c_i x_i$).

We seek to investigate how well the volume formulae of [SL15], comparing the volumes of the polytopes \mathcal{P}_ℓ ($\ell = 1, 2, 3, h$), can be used to predict the quality of the relaxations \mathcal{P}_ℓ ($\ell = 1, 2, 3, h$) as measured by their quasi mean width.

Figure 1 consists of a plot for each scenario: dense, sparse, and very sparse. Each plot illustrates the difference in quasi mean width between \mathcal{P}_3 (using the ‘best’ double McCormick) and each of the other relaxations. Each point represents a choice of bounds and the instances are sorted by $\omega(\mathcal{P}_1) - \omega(\mathcal{P}_h)$. In all three plots $\omega(\mathcal{P}_h) - \omega(\mathcal{P}_3)$ is non-positive, which is to be expected because \mathcal{P}_h is contained in each of the three double-McCormick relaxations. Furthermore the plots illustrate that the general trend is for $\omega(\mathcal{P}_2) - \omega(\mathcal{P}_3)$ and $\omega(\mathcal{P}_1) - \omega(\mathcal{P}_3)$ to be positive and also for $\omega(\mathcal{P}_1) - \omega(\mathcal{P}_3)$ to be greater than $\omega(\mathcal{P}_2) - \omega(\mathcal{P}_3)$. This agrees with Corollary 1 and gives strong validation for the use of volume to measure the strength of different relaxations. It confirms that given a choice of the double-McCormick relaxations, \mathcal{P}_3 is the one to choose.

However, there are a few exceptions to the general trend and these exceptions are most apparent in the very sparse case. In both the dense case and the sparse case we only see a deviation from the trend on a small number of occasions when \mathcal{P}_2 is very slightly better than \mathcal{P}_3 . In each of these cases, the difference seems to be so small that we can really regard \mathcal{P}_3 and \mathcal{P}_2 as being equivalent from a practical viewpoint. In the very sparse case, the general trend is still followed, but we see a few more cases where \mathcal{P}_2 is slightly better than \mathcal{P}_3 . We also see that in a few instances, \mathcal{P}_1 is better than \mathcal{P}_2 and occasionally even slightly better than \mathcal{P}_3 .

However it is important to note that when we consider the sparse and very sparse cases, the differences in quasi mean width between *any* two of the relaxations is much smaller than these differences in the dense case. If we were to take the sparsity of H to the extreme and run our experiments with $n = 60$ and each variable only in one trinomial, the difference in quasi mean width between *any* two of the polytopes will become zero for these boxcup problems. Therefore, it is not surprising that our results diverge from the general trend as H becomes sparser.

Using the common technique of ‘performance profiles’ (see [DM02]), we can illustrate the differences in quasi mean width of the three double-McCormick relaxations in another way. We obtained

the matlab code “perf.m” which was adapted to create these plots from the link contained in [SMM⁺15]. Figure 2 shows a performance profile for each of the dense, sparse and very sparse scenarios. For each choice of bounds, \mathcal{P}_h gives the least quasi mean width (because it is contained in each of the other relaxations). Our performance profiles display the fraction of instances where the quasi mean width of \mathcal{P}_ℓ is within a factor α of the mean width of \mathcal{P}_h , for $\ell = 1, 2, 3$. The plots are natural log plots where the horizontal axis is $\tau := \ln(\alpha)$. Using this measure, we see that the trend in *all* cases is that \mathcal{P}_3 dominates \mathcal{P}_2 which in turn dominates \mathcal{P}_1 . In the very sparse case, we see that \mathcal{P}_3 and \mathcal{P}_2 are very close for small factors α . In general, all three relaxations are within a small factor of the hull. Displaying the results in this manner gives us a way to see quickly which relaxation performs best for the majority of instances. Again, we see agreement with the prediction of Corollary 1 and confirmation that \mathcal{P}_3 is the best double-McCormick relaxation.

3.4 Validating the relationship between volume and mean width

Using the [SL15] formulae, we calculate the volume of the relaxation for each individual trinomial, \mathcal{P}_ℓ . We then we take the fourth root of these volumes and sum over all $|E(H)|$ trinomials to obtain a kind of ‘aggregated idealized radius’ for each relaxation and each set of bounds. Restricting our attention to the dense scenario, in Figure 3, we compare these aggregated idealized radii with quasi mean width, across all relaxations $\mathcal{P}_\ell, \ell = 1, 2, 3, h$ and each set of bounds (each point in each scatter plot corresponds to a choice of bounds). We see a high R^2 coefficient in all cases, so we may conclude that volume really is a good predictor of relaxation width.

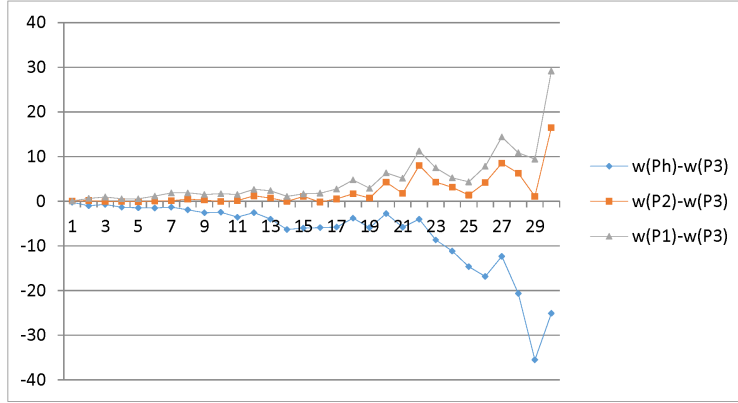
We also compute the difference in width between polytope pairs: \mathcal{P}_h and \mathcal{P}_3 , \mathcal{P}_3 and \mathcal{P}_2 , \mathcal{P}_2 and \mathcal{P}_1 for each direction Q . We then average these width differences for each polytope and each set of bounds, across the 100,000 realizations of Q . We refer to this result as the *quasi mean width difference* of the pair of polytopes. In Figure 4, we similarly compare aggregated idealized radial differences with quasi mean width differences. We see even higher R^2 coefficients, validating volume as an excellent predictor of average objective gap between pairs of relaxations.

3.5 A worst case

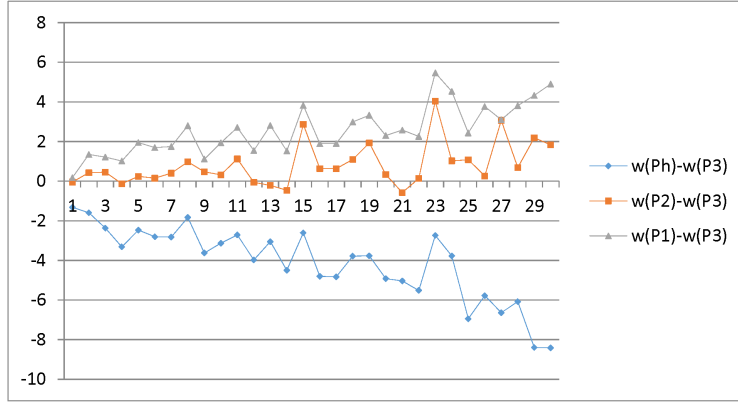
Our final set of experiments relate to a ‘worst case’ as described in [SL15]. In the important special case of $a_1 = a_2 = 0$ and $b_1 = b_2 = 1$, the two ‘bad’ double-McCormick relaxations have the same volume and the ‘good’ double McCormick is exactly the hull. In addition, the greatest difference in volume between the hull and the bad relaxations occurs when $a_3 = \frac{b_3}{3}$.

We compute the same results as we have discussed before (i.e. the differences in quasi mean width between the relaxations) with $n = 6$, but now instead of using random bounds, we fix $a_1 = a_2 = 0$ and $b_1 = b_2 = 1$. We also fix b_3 and run the experiments for $a_3 = 1, 2, \dots, b_3 - 1$. Here, we only consider the $\binom{5}{3} = 10$ trinomials that have the form $x_j x_k x_6$.

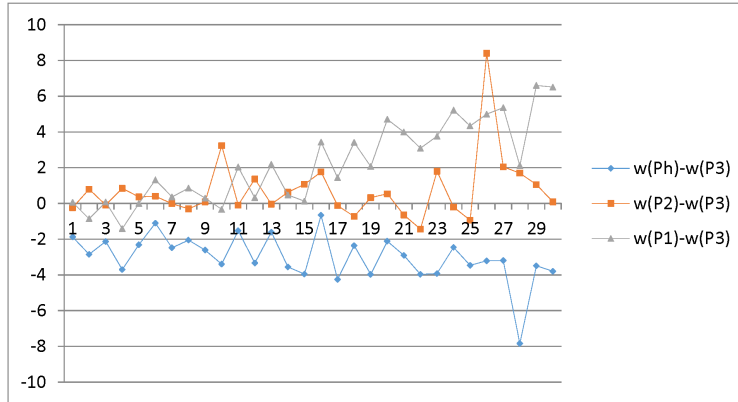
Figure 5 displays a plot of these results for $b_3 = 30, 60, 90, 120$ and 150 . From the inequality systems we know that \mathcal{P}_h is exactly \mathcal{P}_3 , therefore we are interested in the comparison between: \mathcal{P}_2 and \mathcal{P}_3 , and \mathcal{P}_2 and \mathcal{P}_1 . From the plots of these differences, we see exactly what we would expect given the volume formulae. The difference in mean width between \mathcal{P}_2 and \mathcal{P}_1 is very small; from a practical standpoint it is essentially zero. The difference in mean width between \mathcal{P}_2 and \mathcal{P}_3 is always positive, indicating again that \mathcal{P}_3 is the best choice of double-McCormick relaxation. In addition, we observe that the maximum difference falls close to $a_3 = \frac{b_3}{3}$ in all cases, demonstrating again that volume is a good predictor of how well a relaxation behaves.



(a) dense case

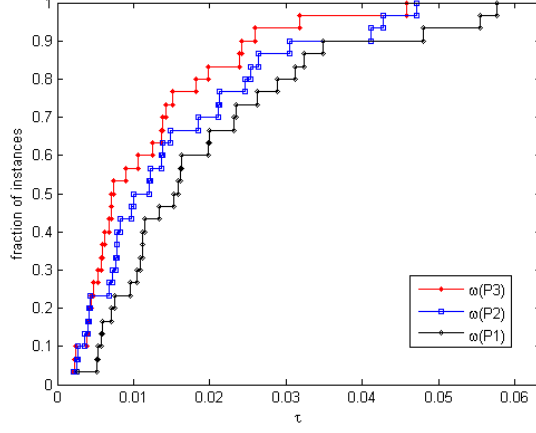


(b) sparse case

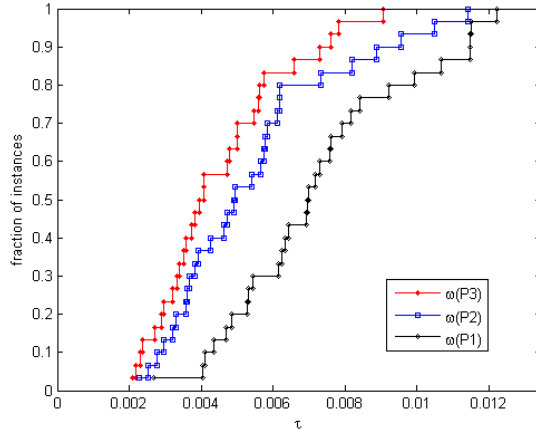


(c) very-sparse case

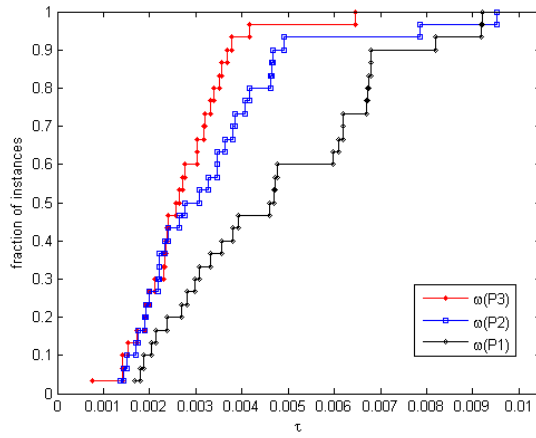
Figure 1: Quasi-mean-width differences



(a) dense case



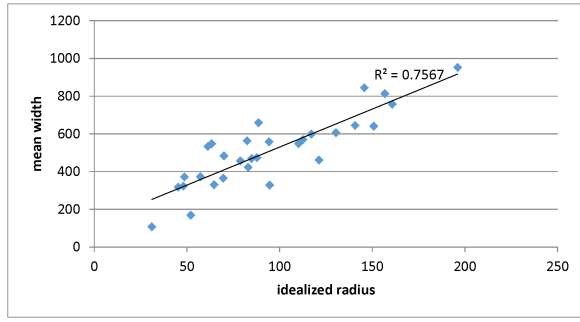
(b) sparse case



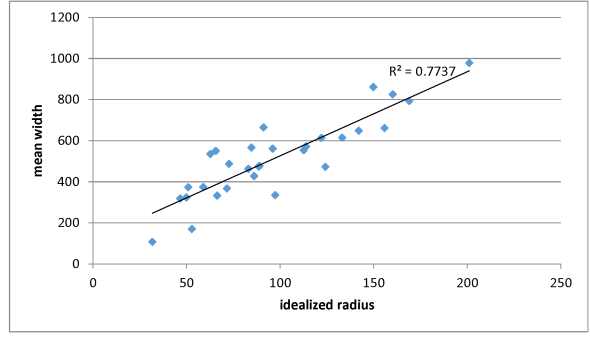
(c) very-sparse case

Figure 2: Quasi-mean-width performance profiles

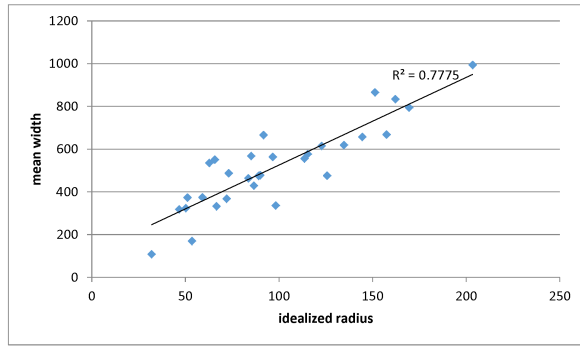
Displays the fraction of instances where the quasi mean width of \mathcal{P}_ℓ is within a factor $\alpha = e^\tau$ of the mean width of \mathcal{P}_h . Note that for small τ , $e^\tau \approx 1 + \tau$.



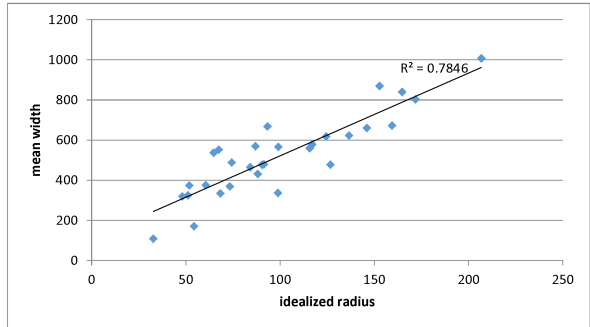
(a) \mathcal{P}_h



(b) \mathcal{P}_3

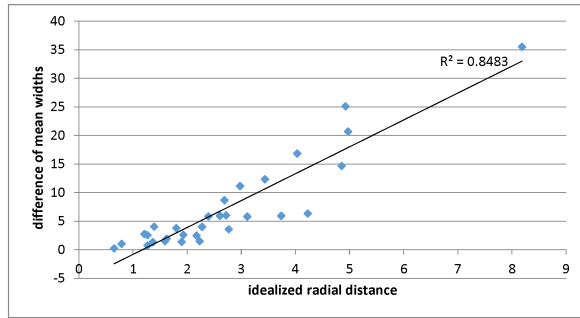


(c) \mathcal{P}_2

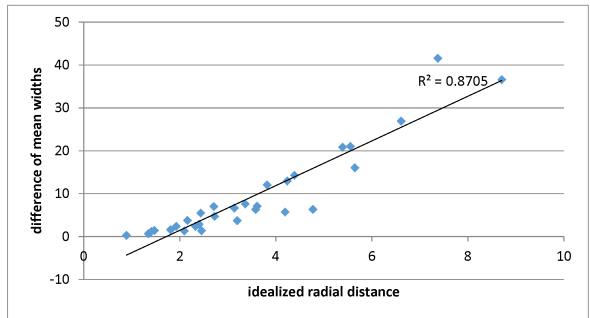


(d) \mathcal{P}_1

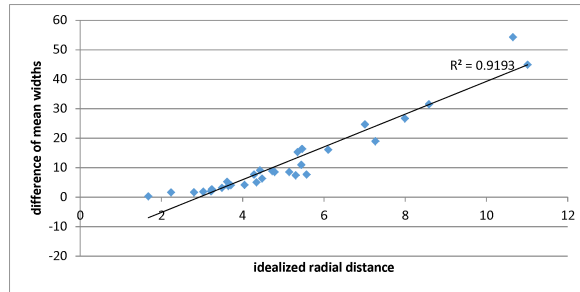
Figure 3: idealized radius predicting quasi mean width



(a) \mathcal{P}_3 vs. \mathcal{P}_h



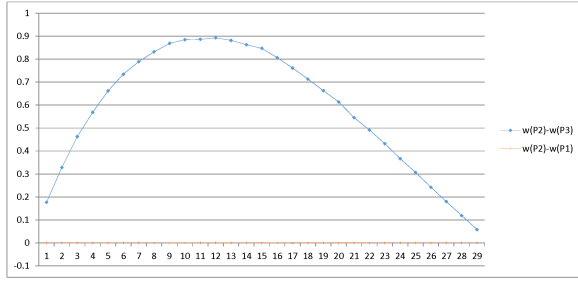
(b) \mathcal{P}_2 vs. \mathcal{P}_h



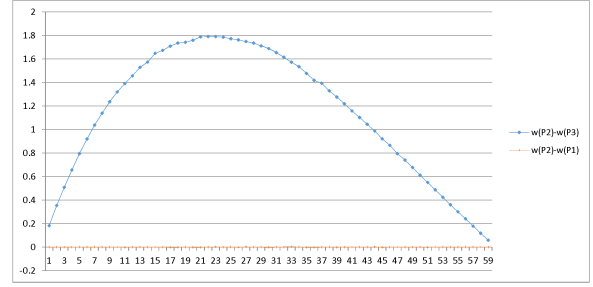
(c) \mathcal{P}_1 vs. \mathcal{P}_h

Figure 4: idealized radial distance predicting quasi mean width difference

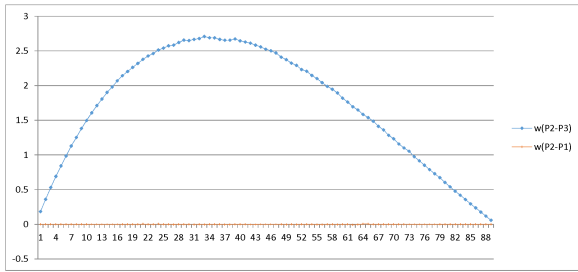
Appendix



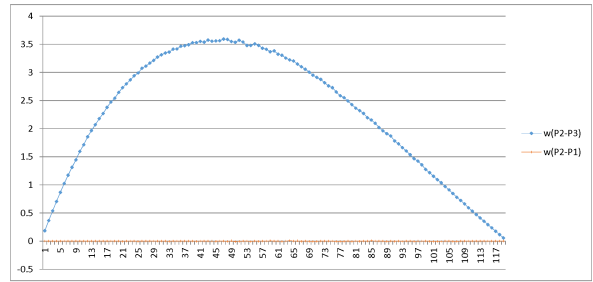
(a) $b_3 = 30$



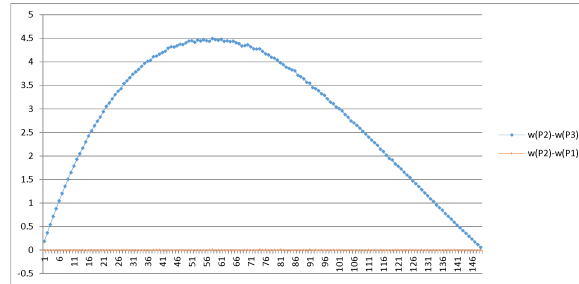
(b) $b_3 = 60$



(c) $b_3 = 90$



(d) $b_3 = 120$



(e) $b_3 = 150$

Figure 5: Worst-case analysis for a_3 ($a_1 = a_2 = 0$, $b_1 = b_2 = 1$)

References

- [ABD10] F. Ardila, C. Benedetti, and J. Doker. Matroid polytopes and their volumes. *Discrete & Computational Geometry*, 43(4):841–854, 2010.
- [ADFN98] C.S. Adjiman, S. Dallwig, C.A. Floudas, and A. Neumaier. A global optimization method, α BB, for general twice-differentiable constrained NLPs: I. Theoretical advances. *Computers & Chemical Engineering*, 22(9):1137–1158, 1998.
- [BDLO13] K. Burggraf, J. De Loera, and M. Omar. *On Volumes of Permutation Polytopes*, pages 55–77. Springer International Publishing, Heidelberg, 2013.
- [BLL⁺09] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wächter. Branching and bounds tightening techniques for non-convex MINLP. *Optimization Methods & Software*, 24(4-5):597–634, 2009.
- [BW91] G. Brightwell and P. Winkler. Counting linear extensions is $\#P$ -complete. In *Proceedings of the Twenty-third Annual ACM Symposium on Theory of Computing*, STOC '91, pages 175–181, New York, NY, USA, 1991. ACM.
- [CLL10] S. Cafieri, J. Lee, and L. Liberti. On convex relaxations of quadrilinear terms. *Journal of Global Optimization*, 47:661–685, 2010.
- [DFK91] M. Dyer, A. Frieze, and R. Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. ACM*, 38(1):1–17, January 1991.
- [DM02] D. Elizabeth Dolan and J. Jorge Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2002.
- [DMW15] Santanu S. Dey, Marco Molinaro, and Qianyi Wang. Approximating polyhedra with sparse inequalities. *Mathematical Programming*, 154(1):329–352, 2015.
- [KLS97] C.-W. Ko, J. Lee, and E. Steingrímsson. The volume of relaxed Boolean-quadric and cut polytopes. *Discrete Mathematics*, 163(1-3):293–298, 1997.
- [Lee07] J. Lee. Mixed integer nonlinear programming: Some modeling and solution issues. *IBM Journal of Research and Development*, 51(3/4):489–497, 2007.
- [LM94] J. Lee and W. Morris. Geometric comparison of combinatorial polytopes. *Discrete Applied Mathematics*, 55:163–182, 1994.
- [McC76] G.P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I. Convex underestimating problems. *Mathematical Programming*, 10:147–175, 1976.
- [MF04a] C.A. Meyer and C.A. Floudas. Trilinear monomials with mixed sign domains: Facets of the convex and concave envelopes. *Journal of Global Optimization*, 29:125–155, 2004.
- [MF04b] C.A. Meyer and C.A. Floudas. Trilinear monomials with positive or negative domains: Facets of the convex and concave envelopes. *Frontiers in Global Optimization*, pages 327–352, 2004.

- [MF14] R. Misener and C. A. Floudas. ANTIGONE: Algorithms for coNTinuous / Integer Global Optimization of Nonlinear Equations. *Journal of Global Optimization*, 2014. DOI: 10.1007/s10898-014-0166-2.
- [Rik97] A. Rikun. A convex envelope formula for multilinear functions. *Journal of Global Optimization*, 10:425–437, 1997.
- [RS96] H.S. Ryoo and N.V. Sahinidis. A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8(2):107–138, 1996.
- [Sah15] N.V. Sahinidis. *BARON 15.6.5: Global Optimization of Mixed-Integer Nonlinear Programs*, User’s Manual, 2015.
- [SL15] E. Speakman and J. Lee. Quantifying double McCormick. arXiv:1508.02966v2, <http://arxiv.org/abs/1508.02966>, 2015.
- [SL16] E. Speakman and J. Lee. On sBB branching for trilinear monomials. GOW ’16 (4 pages) <https://umich.box.com/s/ase34hd2jpb91tk2uqb17ozoytqvhz2c>, 2016.
- [SMM⁺15] A.M. Sofi, M. Mamat, S.Z. Mohid, M.A.H. Ibrahim, and N. Khalid. Performance profile comparison using matlab. In *Proceedings of International Conference on Information Technology & Society 2015*, 2015.
- [SP99] E.M.B. Smith and C.C. Pantelides. A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering*, 23:457–478, 1999.
- [Sta86] R.P. Stanley. Two poset polytopes. *Discrete & Computational Geometry*, 1(1):9–23, 1986.
- [Ste94] E. Steingrímsson. A decomposition of 2-weak vertex-packing polytopes. *Discrete Computational Geometry*, 12(4):465–479, 1994.
- [VG16] S. Vigerske and A. Gleixner. Scip: Global optimization of mixed-integer nonlinear programs in a branch-and-cut framework. Technical Report 16-24, ZIB, Takustr.7, 14195 Berlin, 2016.